An Efficient Hybrid Explicit/Implicit Solvent Method for Biomolecular Simulations

MICHAEL S. LEE,[1,2] FREDDIE R. SALSBURY, JR.,[3] MARK A. OLSON[2]
[1]*CISD, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005*
[2]*Department of Cell Biology and Biochemistry, USAMRIID, 1425 Porter St., Frederick, Maryland 21702*
[3]*Department of Physics, Wake Forest University, Winston-Salem, North Carolina 27109*

Received 14 June 2004; Accepted 13 July 2004
DOI 10.1002/jcc.20119
Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: We present a new hybrid explicit/implicit solvent method for dynamics simulations of macromolecular systems. The method models explicitly the hydration of the solute by either a layer or sphere of water molecules, and the generalized Born (GB) theory is used to treat the bulk continuum solvent outside the explicit simulation volume. To reduce the computational cost, we implemented a multigrid method for evaluating the pairwise electrostatic and GB terms. It is shown that for typical ion and protein simulations our method achieves similar equilibrium and dynamical observables as the conventional particle mesh Ewald (PME) method. Simulation timings are reported, which indicate that the hybrid method is much faster than PME, primarily due to a significant reduction in the number of explicit water molecules required to model hydration effects.

 J Comput Chem 25: 1967–1978, 2004

Key words: molecular dynamics; hydration; generalized Born theory; multigrid; electrostatics

Introduction

In the computer simulation of biomolecules, one often has the choice of treating the aqueous solvent as a collection of explicit water molecules[1,2] or as an implicit solvent model such as the generalized Born (GB) theory.[3] Implicit solvent simulations are typically faster than fully microscopic evaluations and easier to interpret as the water degrees of freedom are absent. Because of their mean-field approximation, implicit solvent models are of lower resolution and have been known to blur the potential energy landscape of a protein,[4] and cause structural distortions.[5] Moreover, implicit models lack local hydrogen bond interactions between the solute and solvent that can lead to incorrect preferences for secondary structural motifs.[6]

Alternatively, explicit solvent methods offer a more detailed and accurate description of a macromolecular system. Yet, these methods often expend a majority of their computational effort on the simulation of water molecules rather than the solute. Furthermore, certain periodic boundary methods, such as particle mesh Ewald (PME), evoke artificial real-virtual solute interactions that can be problematic when determining thermodynamic quantities such as free energies.[7–9]

Methods that combine explicit and implicit solvent approaches[2,10–12] seek to gain the salient features of both approaches; namely, accuracy and speed. There are two classes of hybrid explicit/implicit solvent schemes. In one class, known as reaction field methods,[12–14] the electrostatic interactions between molecules are radially truncated, and a dielectric continuum beyond the cutoff is employed to estimate the dielectric screening of the neglected interactions. An advantage of this method is that periodic boundary conditions can be employed, thus allowing unrestricted movement of water molecules. A further benefit is the use of rather short nonbonded cutoffs (~10 Å). A disadvantage is that the procedure assumes that a dielectric of 80 always lies outside the cutoff region, which is an inappropriate assumption when modeling large solutes such as biomolecules. An additional problem is that the reaction field method uses the same type of periodic box as a conventional Ewald method and thus requires the same number of water molecules to solvate a system.

A second type of hybrid solvent approach, which we will designate as cluster methods,[2,10,15,16] utilizes a simulation volume, often spherical in shape, surrounded by a dielectric continuum.

Correspondence to: M.S. Lee; e-mail: Michael.Lee@amedd.army.mil

Contract/grant sponsor: USAMRIID; contract/grant number: RIID 02-4-1R-069

Contract/grant sponsor: U.S. Army Medical Research and Materiel Command (GETMC task area to M.L. and M.O.)

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
13 JUL 2004	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
An efficient hybrid explicit/implicit solvent method for biomolecular simulations, Journal of Computational Chemistry 25:1967 - 1978	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Lee, MS Salsbury, FR Jr Olson, MA	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD	**RPP-04-314**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

The original document contains color images.

14. ABSTRACT

We present a new hybrid explicit/implicit solvent method for dynamics simulations of macromolecular systems. The method models explicitly the hydration of the solute by either a layer or sphere of water molecules, and the generalized Born (GB) theory is used to treat the bulk continuum solvent outside the explicit simulation volume. To reduce the computational cost, we implemented a multigrid method for evaluating the pairwise electrostatic and GB terms. It is shown that for typical ion and protein simulations our method achieves similar equilibrium and dynamical observables as the conventional particle mesh Ewald (PME) method. Simulation timings are reported, which indicate that the hybrid method is much faster than PME, primarily due to a significant reduction in the number of explicit water molecules required to model hydration effects.

15. SUBJECT TERMS

computer modeling, biomolecular simulation, molecular dynamics, generalized Born theory, multigrid, electrostatics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**12**	

Unlike procedures using periodic boundary conditions (PBC), cluster methods do not allow free movement of water molecules near the simulation surface and thus suffer from surface artifacts[17] that can be partially corrected.[2,10,15,18] Cluster methods are not plagued by the artificial interactions between real and virtual solute atoms found in periodic boundary Ewald methods.[7] Furthermore, cluster methods have been shown to provide convergent observables with increasing simulation radius.[7,10] Unlike reaction field methods, cluster approaches cannot be reliably used with short nonbonded cutoffs; however, long-range electrostatic approximations that greatly reduce computational effort have been devised.[8] The primary savings in finite cluster simulations over conventional periodic boundary Ewald approaches is achieved because fewer numbers of water molecules are required.[2,10]

Still, there are many difficulties with cluster approaches. First, a spherical boundary is often used because there is an analytical solution to the reaction field of a spherical dielectric.[10,19] A spherical shell of water molecules can be, however, quite wasteful for highly nonspherical solutes.[10,20] Simulation volumes that do not have simple geometric shapes require numerical solution of the Poisson equation (or Langevin equation), which is a comparatively slower process.[20] A second problem with cluster approaches when used with an analytical Poisson formula, is that the reaction field potential approaches infinity at the boundary. This is due to the incorrect assumption that individual atoms do not contribute their own dielectric volume as they pierce through the spherical boundary.

There are many cluster approaches that employ irregular simulation volumes that follow the shape of the biological molecule to reduce the number of explicit solvent molecules. For example, Beglov and Roux[21] introduced a method that uses a sum-over-spheres simulation volume and maintains constant pressure[22,23] through dynamic boundaries. Unfortunately, this approach and most other irregularly shaped cluster approaches lack a reaction-field treatment. The one exception is the SAPHYR method of Lounnas et al.,[11] which uses an image dipole formulation to treat the reaction field of each water molecule. Nonetheless, it does not appear that the SAPHYR method incorporates a reaction field for the solute atoms. Perhaps this is because the position of the image dipole is undefined for a solute atom as the atom sits at the center of the sphere it carves out.

In this work, we present a method that also uses a generalized sum-over-spheres boundary but incorporates a continuum reaction field via GB theory. The GB solvent model has been shown to be a reasonable alternative to the solution of the Poisson equation, obtaining an average absolute error of ~1% for a large set of proteins.[24,25] GB solvation energies and forces, unlike Poisson theory, only require a single noniterative solution, and hence are much faster than Poisson methods for arbitrary dielectric volumes. In our approach, the dielectric boundary is held fixed, which further simplifies the problem. With a fixed dielectric boundary, the Born radius becomes a scalar function that can be stored on a grid. This approximation enables atomic Born radii to be obtained expediently by interpolation.

With the time-consuming Born radii integration step removed, the bottlenecks in our method become the pairwise electrostatic and GB terms. Use of a simple nonbonded cutoff in explicit water simulations is well known to degrade accuracy.[7] Approximate treatments of the long-range electrostatic component include the extended electrostatics approach,[26] the local reaction field method,[8] and the fast multipole method.[27] Nonetheless, it is not obvious how to modify these methods to handle GB terms because they are all based on mathematical manipulations of the Coulombic kernel. Fortunately, the multigrid approach[28,29] is an alternative strategy for reducing computational cost for pairwise interactions that does not require a Coulombic kernel. The multigrid approximation to pairwise point charge interactions has linear scaling computational behavior with respect to system size and early crossover with respect to exact calculations.[29] Furthermore, the accuracy of the multigrid approach is tunable and within the error tolerances required for molecular dynamics simulations.[29] In this work, we modified the multigrid approach of Skeel et al.[29] to approximate the long-range portion of the pairwise GB term.

In the remainder of this work, we present the details of our hybrid solvent method and multigrid algorithm. Then, we present results showing the computational benefits of the multigrid procedure over a standard cutoff approach. Next, we compare dynamical and structural results[15,30] of several model simulations between our procedure and the conventional PME explicit solvent method. Finally, we envisage the types of applications where the hybrid method may have an advantage over conventional PBC approaches.

Theory and Methods

Definition of Hybrid Solvent Potential

The goal of a hybrid solvent approach is to partition solvent degrees of freedom into explicit and implicit regions.[10,31,32] In our case, a simulation volume is defined as a region in space that contains the solute and explicit water molecules. Outside the simulation volume, an implicit solvent is described by a high dielectric continuum. The dielectric continuum can be thought of as thermal reservoir of infinitesimal dipoles that interact with the solute, explicit water atomic charges, and themselves. Figure 1 schematically illustrates the system setup for a hybrid-layer simulation.

Previous reaction field approaches have often been limited to spherical volumes, as one of the few analytic solutions to the Poisson equation is for a spherical boundary.[10] GB theory offers a fast, yet approximate, analytical treatment of the dielectric continuum that is not limited to a spherical boundary. However, we have chosen a rather simple dielectric volume that is composed of a transformed sum-over-spheres. Combining spheres of radius w around each heavy atom is nearly identical to describing a layer of width, w, surrounding the entire solute. To assure continuity at the intersection of spheres centered on different atoms, we use a volume function similar to the one proposed by Im et al.[33] If we define a volume function, $V(\vec{r})$, such that $V = 1$ is defined as inside and $V = 0$ is defined as outside, the interlocking-sphere volume has the following form:

$$V(\vec{r}) = 1 - \prod_i \left[1 - v\left(\frac{|\vec{r} - \vec{x}_i| - w_i^0}{w_i^1 - w_i^0} \right) \right], \qquad (1)$$

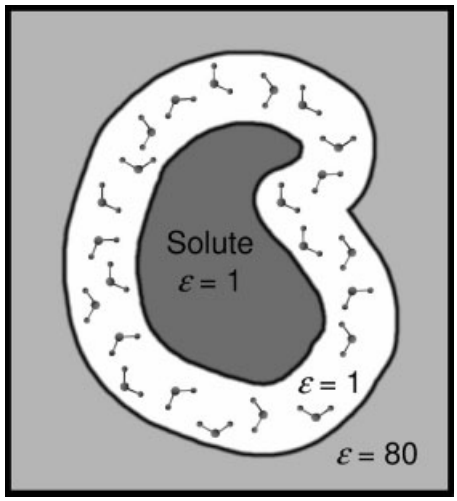

Figure 1. Schematic of the hybrid explicit/implicit solvent approach. The simulation volume consists of a solute immersed in a layer of water molecules. The gray outer region corresponds to the bulk dielectric continuum.

where

$$v(s) = \begin{cases} 1 & s \le 0 \\ 1 - 6s^5 + 15s^4 - 10s^3 & 0 < s < 1, \\ 0 & s \ge 1 \end{cases} \qquad (2)$$

$\vec{x}_i$ is the position of the center i, and w_i^0 and w_i^1 are the starting and stopping points, respectively, of the atomic function tail. Given this volume function, it is now possible to define the reaction field and repulsive boundary potential; however, to increase flexibility in parameterization, the values of w_i^0 and w_i^1 are different for the two energy terms (see below).

The reaction field is defined using GB theory. Because the simulation and dielectric volumes are assumed fixed, Born radii are precalculated on a three-dimensional grid with a cell size of 2 Å. During the simulation, the Born radius for each atom is obtained by cubic interpolation (see below) of the Born radii grid. The Born radii, α_m, of each grid point, m is defined in terms of the volume function, $V(\vec{r})$:

$$\alpha_m^{-1} = \left(1 - \frac{1}{\sqrt{2}}\right)\left(R^{-1} - \frac{1}{4\pi}\int\int_R^\infty \frac{V(\vec{r})}{|\vec{r} - \vec{x}_m|^4}\, dr d\Omega\right) + \left(\frac{1}{4R^4} - \frac{1}{4\pi}\int\int_R^\infty \frac{V(\vec{r})}{|\vec{r} - \vec{x}_m|^7}\, dr d\Omega\right)^{1/4}, \qquad (3)$$

where R is an asymptotic radius set to 1.5 Å, and $\vec{x}_m$ is the position of grid point, m. After α_m is calculated via eq. (3) it is linearly transformed, $\alpha'_m = 1.0649\alpha_m - 0.316$. This transformation was derived empirically by fitting the GB self energies of a 1000 random points inside of a sphere to the analytical Poisson results.[10] Equation (3) is evaluated via a numerical integration algorithm that uses a 26-point Lebedev angular grid and 17 radial points.[25] Furthermore, eq. (3) has been shown elsewhere[25,34] to provide an accurate estimate of Born radii compared to the Poisson method, with correlation coefficients of 0.99.

Given Born radii, the total GB energy is calculated via a formula similar to the one originally proposed by Still et al.,[3]

$$E_{GB} = k\left(1 - \frac{1}{\varepsilon_{solv}}\right)\sum_{ij}\frac{q_i q_j}{r_{ij} + \frac{1}{2}(\alpha_i + \alpha_j)\exp[-2r_{ij}/(\alpha_i + \alpha_j)]}, \qquad (4)$$

where q_i are atomic charges and r_{ij} is distance between atoms i and j, ε_{solv} is the dielectric constant of the solvent (for this work, $\varepsilon_{solv} = 80$) and $k = -166.0$. We used this alternative formula because it appeared to provide better agreement with the exact results of a spherical system (results not shown). Also, eq. (4) reduces the number of square root evaluations by a factor of 2 vs. the original formula. Nonetheless, a variety of alternative GB formulas are possible that provide similar results as the original Still formula.[35,36] The wall potential in our formulation is also defined in terms of the volume function [eq. (1)]:

$$E_{wall}(\vec{r}) = \frac{1}{E_{wall}^{max^{-1}} + V(\vec{r})} - \frac{1}{E_{wall}^{max^{-1}} + 1}, \qquad (5)$$

such that $V = 1$ maps to $E_{wall}(\vec{r}) = 0$ and $V = 0$ maps to $E_{wall}(\vec{r}) \approx E_{wall}^{max}$. Because we use a finite value for E_{wall}^{max}, the wall potential is actually penetrable. This turns out to be useful because a system that is started with too many water molecules will actually spill the excess water molecules beyond the simulation volume, thus relieving some pressure on the system. We did not incorporate an attractive van der Waals (vdW) potential because initial tests indicated that such a potential had a relatively small impact on radial distribution functions. A more important factor was the proper placement of the dielectric boundary for the reaction field. We also did not incorporate energy terms that would improve the angular isotropy of water molecules near the cluster surface.[2,18] Such correction terms are supposed to reduce artificial forces that arise from surface dipoles. However, one key study indicated that these terms fall short in reducing surface dipole artifacts in the application of calculating ionic charging free energies in solution.[17] Furthermore, implementing angular correction terms for nonspherical boundaries is nontrivial and beyond the scope of this work.

Although we present the hybrid method as a means to incorporate layer of explicit water molecules around a solute, it can also be used to simulate a spherical cluster of water molecules. First, if a solute is present, its center of mass is translated to the origin. Next, a fixed dummy atom is defined at the origin. This dummy atom is the only atom used to define the simulation volume. The simulation sphere radius, R_{sphere}, is calculated as a sum of the specified layer width w and the maximal distance between a solute atom and the origin. This simple protocol could be improved, because it does not provide the optimally smallest simulation

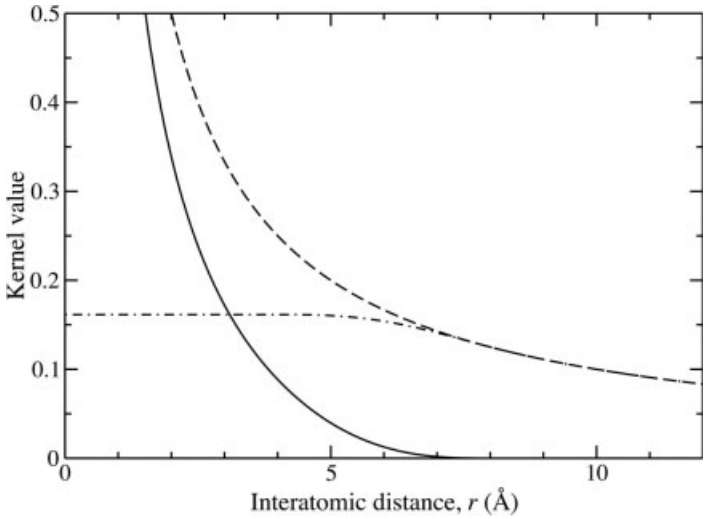

Figure 2. The interaction kernels used in the multigrid implementation of the Coulomb electrostatic term. In this example, the cutoff value, a, is 8 Å. Dashed line: Coulomb kernel, $1/r$, solid line: local kernel, $K_{\text{local}}(r)$, dot-dashed line: soft kernel, $K_{\text{soft}}(r)$.

sphere that can encompass the solute and still provide at minimum a width, w, of water molecules.[37]

Multigrid of Pairwise Interactions

Having simplified the normally expensive Born radii calculation, the Coulomb and GB pairwise interactions become the most computationally expensive terms in the hybrid solvent method. Thus, we implemented a multigrid approach to approximate the long-range component of the electrostatic and GB interactions, and hence greatly reduce the computational cost. The multigrid method for Coulomb electrostatics has been developed by Skeel et al.[29] and Sandak.[28] We adapted the multigrid method of Skeel et al.[29] to incorporate the GB pairwise screening formula [eq. (4)].

The basic principle of a multigrid electrostatics procedure is to cast pairwise charge interactions onto grids through interpolation and treat various interaction distances with different degrees of approximation or grid resolutions (indexed as L). The multigrid algorithm is sketched as follows:

1. Perform explicit local interactions of atomic charges, when $r <$ a, where a is user-defined local cutoff ($L = 0$).
2. Interpolate atomic charges onto the finest resolution charge grid ($L = 0 \rightarrow L = 1$).
3. Create a hierarchy of grids by interpolating each charge grid to the charge grid with the next coarsest resolution ($L = n \rightarrow L = n + 1$).
4. For each individual charge grid, L, build a potential grid such that each grid point sees only charges less than a distance $2^L a$ away.
5. Interpolate potential grids from coarse to fine resolution ($L = n + 1 \rightarrow L = n$).
6. Contract atomic charges with the finest resolution potential grid and combine with the local interactions in step 1 to obtain the energy and forces ($L = 1 \rightarrow L = 0$).

To separate the ranges in step 4 precisely, the interaction kernel for each grid level L, $K^L(r)$, is split into two components: a soft function and a local function,

$$K^L(r) = K^L_{\text{soft}}(r) + K^L_{\text{local}}(r). \tag{6}$$

The soft function for the $L = 0$ interactions must slowly vary because it is the basis of interaction for the $L = 1$ grid. For a Coulombic potential, $K^0(r) = 1/r$, the soft function is chosen such that $K^0_{\text{local}}(r)$ goes to zero beyond cutoff a. In the original article of Skeel et al.,[29] several soft functions are presented and tested. Nonetheless, we desired to reduce the variance of the soft function completely in the domain, $r < a/2$, so that self-interaction and exclusion interaction artifacts could be reduced or eliminated. Therefore, we created a new piecewise soft potential that is continuous up to second derivatives at $r = a/2$ and $r = a$;

$$K^0_{\text{soft}}(r) = \begin{bmatrix} \dfrac{6}{a^5}\left(r - \dfrac{a}{2}\right)^4 - \dfrac{16}{3a^4}\left(r - \dfrac{a}{2}\right)^3 + \dfrac{31}{24a}, & \dfrac{a}{2} < r < a \\ \dfrac{31}{24a}, & r < \dfrac{a}{2} \end{bmatrix}. \tag{7}$$

The local function for $L = 0$, by definition, is the difference of the original kernel and the soft function,

$$K^0_{\text{local}}(r) = \frac{1}{r} - K^0_{\text{soft}}(r). \tag{8}$$

The interaction kernels for $L = 0$ are depicted in Figure 2. Analogously, for each grid level, L, except the coarsest, $L = L_{\max}$, the interaction kernel is composed of a soft function where a in eq. (7) is replaced by $2^{L-1}\,a$, minus a softer function, where a is substituted with $2^{L}a$. For the coarsest grid, $L_{\max}$, the interaction kernel is simply the softest function corresponding to a in eq. (7) being replaced by $2^{L_{\max}-1}\,a$. In this work, $L_{\max}$ was set to 3. Larger values of $L_{\max}$ might be required for system sizes larger than the ones treated here.

Multigrid steps 2, 3, 5, and 6 involving cubic interpolation utilize six-dimensional basis functions of the form:[29]

$$\phi_L(x_a, y_a, z_a; x_b, y_b, z_b) = \Phi\left(\frac{x_a - x_b}{h_L}\right)\Phi\left(\frac{y_a - y_b}{h_L}\right)\Phi\left(\frac{z_a - z_b}{h_L}\right), \tag{9}$$

where

$$\Phi(p) = \begin{cases} (1 - |p|)\left(1 + |p| - \frac{3}{2}p^2\right) & |p| \le 1 \\ -\frac{1}{2}(|p| - 1)(2 - |p|)^2 & 1 \le |p| \le 2 \\ 0 & |p| \ge 2 \end{cases}, \tag{10}$$

where (x_a, y_a, z_a) and (x_b, y_b, z_b) are 3D coordinates corresponding to the source, a, and destination, b, of a charge or potential, L is the level of the destination grid, and h_L is the side length of a destination grid cube. For example, in step 2, (x_a, y_a, z_a) corresponds to the atomic coordinates, and (x_b, y_b, z_b) refers to the $L =$ grid coordinates. In step 3, (x_a, y_a, z_a) signifies the $L = n$ grid coordinates and (x_b, y_b, z_b) corresponds to the $L = n + 1$ grid coordinates. In this work, the highest resolution grid was assigned a cell size, $h_L = 4$ Å. This value was chosen based on a compromise between speed and accuracy. A detailed analysis using various values of h_L, several interpolation formulae, and alternative soft functions is presented in Skeel et al.[29]

The multigrid procedure is one of the simplest algorithms for reducing the computational cost of long-range electrostatic interactions; however, certain artifacts associated with casting an atomic problem onto a grid are to be anticipated. For instance, the multigrid algorithm erroneously includes the excluded self, 1–2, 1–3, and 1–4 Coulomb interactions. These interactions are approximately subtracted out from the energy by simply including the explicit short-range terms without the Coulomb kernel term [eq. (8)]. Nevertheless, grid artifacts remain unless other strategies are employed.[28,29] For example, in a test calculation that we conducted, the Coulombic self-energy of an isolated sodium ion ranged from 1.5 to 3.8 kcal/mol, depending on its position relative to the grid. We did not choose to remove grid artifacts in this work; nevertheless, understanding the effects of these artifacts and removing them should be the topic of future work.

To extend the multigrid procedure to handle the GB solvent model, the kernel from eq. (4),

$$J = \left\{ r_{ij} + \frac{1}{2}(\alpha_i + \alpha_j)\exp[-2r_{ij}/(\alpha_i + \alpha_j)] \right\}^{-1}$$

is also split into local and soft terms, J_{local} and J_{soft}:

$$J^L(r_{ij}, \alpha_i, \alpha_j) = J^L_{\text{local}}(r_{ij}, \alpha_i, \alpha_j) + J^L_{\text{soft}}(r_{ij}, \alpha_i, \alpha_j). \tag{11}$$

Some experimentation determined that a reasonable splitting involves setting r_{ij} equal to the corresponding inverse of the respective Coulomb terms, K^L_{local} and K^L_{soft}:

$$\begin{aligned} J^L_{\text{local}}(r_{ij}, \alpha_i, \alpha_j) &= J([K^L_{\text{local}}]^{-1}, \alpha_i, \alpha_j) \\ J^L_{\text{soft}}(r_{ij}, \alpha_i, \alpha_j) &= J([K^L_{\text{soft}}]^{-1}, \alpha_i, \alpha_j). \end{aligned} \tag{12}$$

Furthermore, Born radii are evaluated at every multigrid point. Born radii at the lowest level multigrid ($L = 1$) are obtained via bilinear interpolation of the same Born radii grid used to determine atomic Born radii. Coarser grids ($L > 1$) are obtained by cubic interpolation of their next finest resolution Born radii grid ($L - 1$).

Hybrid Simulation Protocol

The hybrid solvent method described in this work was incorporated into the CHARMM program (version c30b1).[38] Molecular dynamics simulations in this work were performed with the PARAM22 empirical forcefield potential.[39] The equations of motion were integrated with a Langevin dynamics algorithm at constant temperature (298 K) using a Langevin friction constant of 1 ps^{-1} and an integration time step of 2 fs. Covalent bonds between heavy atoms and hydrogens were constrained using the SHAKE algorithm.[40] In the dynamics simulations, the local cutoff for the multigrid procedure, a, was set to 8 Å. The vdW terms were truncated with a simple cutoff that switched off from 7 to 8 Å. Pure solvent and single ion simulations (Na^+, Cl^-) were run for a total simulation time of 1 ns.

Two hybrid-based molecular dynamics simulations of a protein were performed in this work. One simulation involved a 10-Å layer ($w = 10$ Å) of water molecules surrounding the protein. The other simulation involved a spherical cluster of water molecules with at least 10-Å coverage for every atom on the protein surface. We also ran a layer calculation that excluded the entire GB component, but otherwise was identical to the hybrid layer simulation. The choice of a 10-Å layer was based on the observation that charging free energies of ions approximately converge at that size (results shown below).[7,41] Production simulations were run for 2 ns with structures saved every picosecond. The protein used in this work was the 62-residue B1 domain of protein L (PDB: 1HZ6).[42] The protein and water molecules were first minimized for 100 steps using the adopted-basis Newton–Raphson method to remove bad vdW contacts. This was followed by allowing the water molecules to equilibrate for 10 ps with the solute heavy atoms held fixed.

The general procedure for preparing a hybrid solvent calculation consists of two steps: (1) determining of the number of water molecules needed to fill the simulation volume, and (2) carving out these water molecules in the shape of a layer or sphere from a large block. The target number of water molecules is calculated as a product of the water simulation volume (WSV) and a specified bulk water density of 0.0334 $Å^{-3}$. The WSV is equal to the difference between the simulation volume and the standard solute

Table 1. Accuracy of Multigrid (MG) and Standard Cutoff (STD) Approaches for the Model System of a Sodium Ion Embedded in a 17-Å Sphere of Water Molecules.

Method	Cutoff start (Å)	Cutoff stop (Å)	*% EE*	*% AFE*	*% MFE*
MG	8	8	0.24	3.4	12
MG	10	10	0.053	1.7	8.7
MG	12	12	0.014	0.97	3.9
STD	8	10	0.13	49	220
STD	10	12	0.50	35	160
STD	14	16	3.0	19	94
STD	18	20	0.75	8.4	36
STD	22	24	0.51	3.1	15
STD	26	28	0.0018	0.78	4.4

Definitions of error measures are described in the text. Energies and forces consist of the electrostatic and GB terms only.

vdW volume. For the layer simulations, the simulation space is obtained by setting all of the heavy atomic radii to w. The *coor volume* command in CHARMM was used to numerically calculate the simulation volume and the solute vdW volume. For the spherical boundary simulations, the simulation volume is equal to $4\pi/3\ R^3_{\text{sphere}}$. With a target number of water molecules determined, a large cubic box of water molecules is overlaid onto the solute. First, the water molecules that are less than 3.1 Å from a solute heavy atom are deleted. Next, for the layer approach, an iterative procedure is performed in which water molecules beyond a certain cutoff distance from all heavy atoms of the solute are deleted. The procedure repeats deletion with decreasing cutoffs until the number of water molecules remaining is less than or equal to the target number of water molecules. The spherical simulations utilize instead a slowly decreasing radius. Often, in actual system setups, some water molecules will be initially positioned outside the wall potential. Due to the finite nature of the wall potential, these water molecules will drift away during a simulation. Thus, we use the MMFP module in CHARMM to place a large spherical boundary potential around the simulation volume to prevent the unbounded drift of water molecules that initially escape from the smaller finite boundary of eq. (5).

The placements of the dielectric boundary and wall potential were empirically parameterized to achieve a reasonable radial distribution function for a spherical droplet of bulk water with a radius of w. The wall potential was placed between $w_0(\text{wall}) = w - 0.6$ Å and $w_1\ (\text{wall}) = w + 0.4$ Å with a height, $E^{\max}_{\text{wall}}$, arbitrarily set to 30 kcal/mol. The dielectric boundary, which should roughly correspond to the average vdW volume of the bounded explicit waters, ranged from $w_0(\text{dielectric}) = w + 1.8$ Å to $w_1(\text{dielectric}) = w + 2.2$ Å.

Solute Restraints

Because the simulation volume in our method is fixed in space, it is necessary to restrain solute translations in both the sphere and layer methods and solute rotations in the layer method. The translational restraint on the center of geometry (COG) of the solute has the form:

$$E_{\text{COG}} = k_{\text{COG}} \left| \left(\frac{1}{N} \sum_i \vec{x}_i \right) - \vec{x}_0 \right|^2, \tag{13}$$

where N is the number of solute atoms, $\vec{x}_0$ is the origin of the simulation volume, and k_{COG} is set to 100 kcal/mol/Å^2. The rotational restraint (RR) on the solute involves restricting rotation around each coordinate axis. For example, the z-axis restraint has the form:

$$E^{xy}_{RR} = \frac{1}{2} k_{RR} \left[\frac{1}{2} \arctan\left(\frac{-2I_{xy}}{I_{xx} - I_{yy}} \right) - \theta_0 \right]^2, \tag{14}$$

where $\mathbf{I}$ is the moment of inertia tensor and θ_0 is the reference angle, which is equal to zero when the solute is initially transformed to standard orientation. For the layer simulations in this work, k_{RR} is set to 100 kcal/mol/radians^2.

Periodic Boundary Simulation Protocol

The periodic boundary condition simulations in the work were performed with the PME method[43] as implemented in CHARMM. Local electrostatic and vdW interactions were truncated with a switching function from 7 to 8 Å. In the protein simulation, a cubic volume with at least 10 Å of space between the initial conformation of the solute and the wall was employed leading to dimensions of 58 × 58 × 58 Å, and a total of 6231 water molecules. The 3D fast Fourier transform grid had dimensions of 64 × 64 × 64. Langevin dynamics simulations with a 1 ps^{-1} friction constant were run with an extended system constant pressure algorithm using a reference pressure of 1 atm, a piston mass of 100 amu and a collision frequency of 10 ps^{-1}.

Results and Discussion

In Tables 1 and 2, the accuracy of the multigrid (MG) approximation is compared to the standard switching cutoff (STD) approach for two model systems: a sodium ion embedded in a 17-Å sphere of TIP3 water molecules, and two proteins surrounded by a 10-Å layer of water molecules. Three accuracy measures were used to evaluate the multigrid and cutoff approximations to the combined electrostatic and GB terms:[29] energy error (EE),

Table 2. Accuracy of Multigrid (MG) and Standard Cutoff (STD) Approaches for Two Proteins Embedded in a 10-Å Solvent Layer: (a) Turkey Ovomucoid Receptor (PDB Identifier: 1OMT),[54] (b) Trypsin (PDB Identifier: 1TNJ).[55]

Method	Cutoff start (Å)	Cutoff stop (Å)	% *EE*	% *AFE*	% *MFE*	Time (s)
(a)						
MG	8	8	0.55	4.5	19	36
STD	26	28	0.28	3.8	23	224
(b)						
MG	8	8	0.25	5.2	27.3	88
STD	26	28	0.095	8.0	55.8	993

Definitions of error measures are described in the text. Energies and forces consist of the electrostatic and GB terms only. Last column indicates CPU time necessary to run 100 steps of geometry optimization.

$$\%EE = 100\% * \frac{|E - E_{\text{exact}}|}{|E_{\text{exact}}|}, \tag{15}$$

average force error (AFE),

$$\%AFE = 100\% * \frac{\sum_i m_i^{-1/2}|F_i^{\text{exact}} - F_i|}{\sum_i m_i^{-1/2}|F_i^{\text{exact}}|}, \tag{16}$$

and maximum force error (MFE),

$$\%MFE = 100\% * \frac{\max[m_i^{-1/2}|F_i^{\text{exact}} - F_i|]}{N^{-1}\sum_i m_i^{-1/2}|F_i^{\text{exact}}|}, \tag{17}$$

where m_i is the mass of atom i, E is the electrostatic + GB energy, F_i is the total electrostatic + GB force acting on atom i, the superscript "exact" corresponds to the infinite cutoff result, and N is the total number of atoms. As can been seen in Tables 1 and 2, a cutoff larger than 20 Å is required to match the results of the multigrid approach that uses a local cutoff of 8 Å. Table 1 also suggests that force errors in the standard approach are still significant even when the cutoff exceeds the simulation radius such that all sodium–water interactions are included. In comparison, the multigrid approach does a reasonable job in adding back the water dipole–dipole interactions even though the highest resolution multigrid cell is larger than a water molecule. Furthermore, in Table 2, a comparison of the CPU times require to perform 100 steps of minimization for two methods that have similar accuracy (MG 8 and STD 26/28) suggests that the multigrid method is approximately 10 times faster than the standard cutoff approach. All CPU times stated in this work were obtained from simulations run on an AMD Athlon™ MP 2200+ computer using the Portland Group Fortran compiler (version 4.0). Finally, in Figure 3, note that the multigrid technique has linear scaling behavior with respect to system size for a series of spherically solvated Na^+ systems of increasing size. It is interesting that the standard procedure with large cutoffs actually does not show linear scaling behavior until ~7000 atoms. Another issue is how accuracy is affected by increasing the system size. As can be seen in Figure 3, the error for multigrid is relatively flat vs. system size. The fact that it does not perform better with small system sizes is a probably a result of the short exact evaluation cutoff (8 Å) and the grid artifacts mentioned above.

Next, we evaluate some of the structural properties of a spherically constrained bulk water system using the hybrid method. In Figure 4, one can see that the hybrid method provides a reasonably homogeneous radial distribution of oxygen centers for three dif-

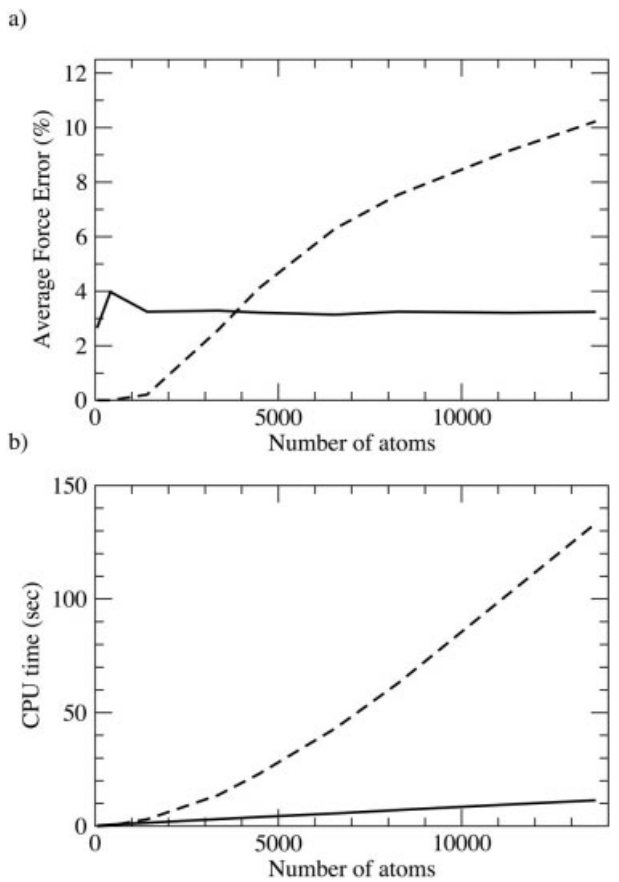

Figure 3. Comparison of multigrid (a = 8 Å) vs. standard cutoff (26/28 Å) for spherical solvated Na^+ systems of various sizes. (a) Average force error and (b) CPU time. Solid line: multigrid, dashed line: 26/28 Å switched cutoff.

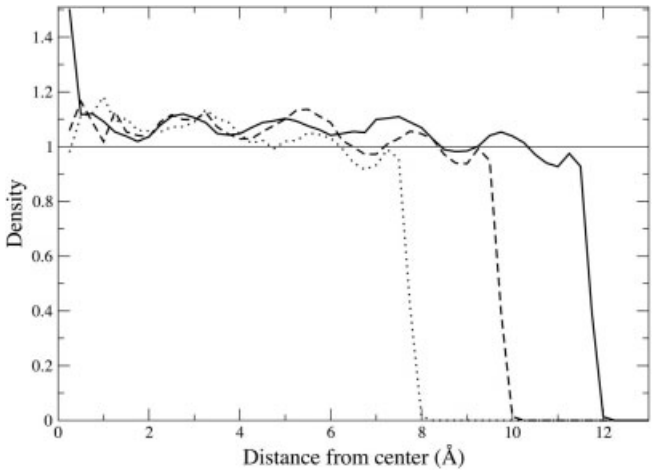

Figure 4. Radial density profiles for spherical clusters of pure solvent. Solid line: $w = 12$ Å, dashed line: $w = 10$ Å, dotted line: $w = 8$ Å.

ferent system sizes. However, two notable deviations from homogeneity exist. First, oscillations with a magnitude of ~0.05 and periodicity of ~2–3 Å exist in all three curves.[10] Likely, this is due to edge effects that propagate from the boundary. The other deviation from homogeneity is the systematic decline of density from the center to the boundary. This effect is probably the result of a moderate imbalance between the intramolecular and boundary forces. A slight reduction in the density of the system reduces the oscillatory behavior; however, the systematic density decline becomes more pronounced.

Figure 5 illustrates that the angular distribution of the water molecules near the surface deviate considerably from ideal bulk isotropic behavior, which is a sine function, (1/2) sin(θ). Most importantly, the dipole distribution graph is significantly skewed

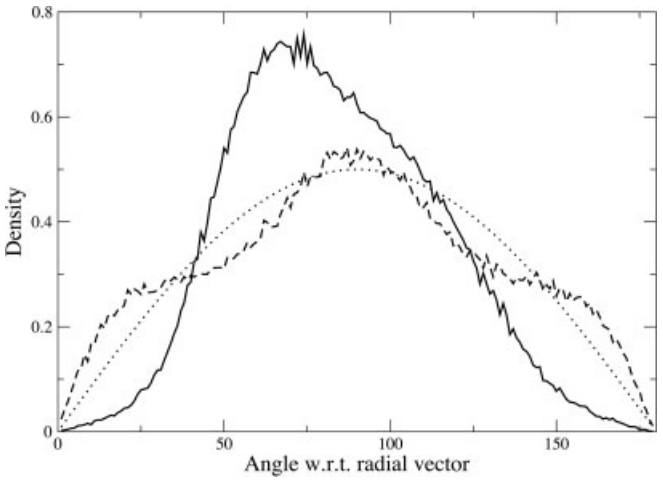

Figure 5. Angular distributions of water molecules that are 1 Å or less from the surface boundary for a spherical cluster of pure solvent ($w =$ 12 Å). Solid line: distribution of dipole vectors, dashed line: distribution of normals to the plane of water molecule, dotted line: ideal curve for both distributions, 1/2 sin (θ).

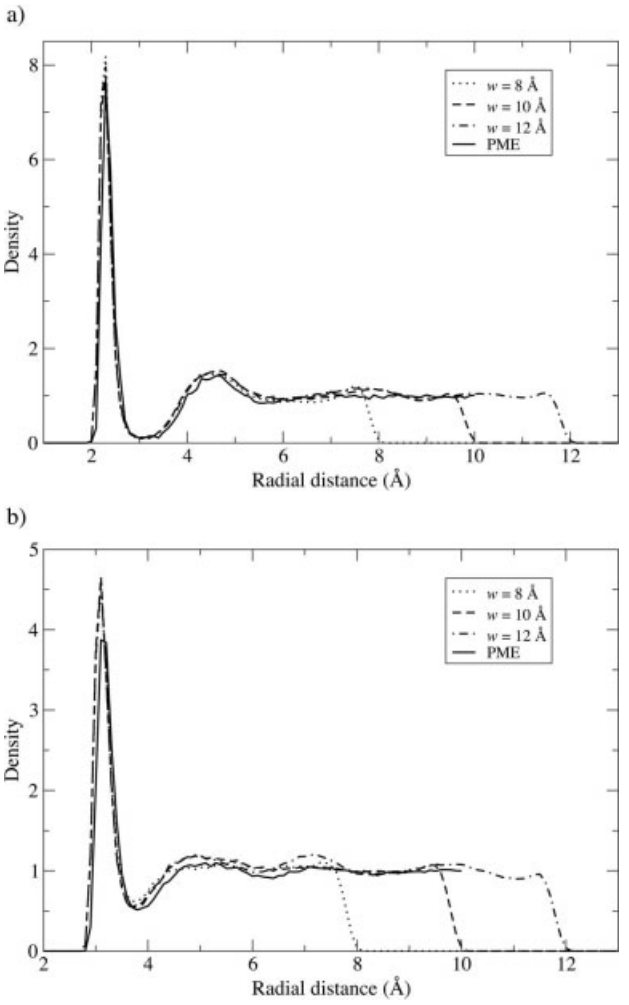

Figure 6. Radial distribution functions of oxygen atom for (a) sodium and (b) chloride ions in water. Lines labeled "$w = \ldots$" correspond to sperical cluster calculations.

towards 75 degrees, which roughly corresponds to surface water molecules with one hydrogen atom pointing outwards from the surface. The distribution of the normal perpendicular to the plane is somewhat better, although significant deviation from isotropy exists. We tested angular correction terms for spherical boundaries similar to those found in other works[10,18] and found improved isotropy for both angular distributions (results not shown). However, we did not implement the correction terms for this work, for two reasons. First, we found that despite improving the angular distributions, the dipolar artifact near the boundary, which can lead to spurious forces on the solute, did not significantly diminish.[44] Second, we were unsure that applying an empirical angular correction derived from spherical calculations would generally be applicable to a layer where variations in curvature exist. Perhaps the angular distribution protocol in the SCAAS method[2] could be incorporated, where the desired distribution could be imposed through angular restraints.

Figure 6 shows the calculated radial distributions of the water oxygen atoms for two simple monatomic solute systems, Na^+ and

Table 3. Charging Free Energies (kcal/mol) of Atomic Ions, Na^+ and Cl^-, in Spherical Clusters of Different Sizes.

Radius (Å)	$Na^{+\,a}$	$Na^{+\,b}$	$Cl^{-\,a}$	$Cl^{-\,b}$
8	−105.0	−105.2	−82.6	−82.4
10	−105.7	−106.0	−82.9	−84.2
12	−106.0	−106.2	−83.7	−83.8

Charging free energies were determined by thermodynamic integration[56] over 10 equally spaced windows with charging parameters, $\lambda = 0.05, 0.15, \ldots, 0.95$. Each window consisted of 20 ps of equilibration and 80 ps of production. Integration was performed via a simple linear fit of the energy derivatives. Convergence errors are estimated to be ~0.5 kcal/mol based on a few repeated simulations that were started with different random seeds.

[a]Multigrid algorithm was used with local cutoff $a = 8$ Å.

[b]No cutoffs were used.

Cl^- ions. Simulations of spheres with radii of 8, 10, and 12 Å were compared with PME calculations using a box size of 18 Å. The spherical clusters are in good agreement with PME with respect to the locations of the first two peaks, although the cluster methods tend to overestimate the size of the peaks. This problem is likely related to the results in Figure 4, which suggest that the water density is slightly overestimated near the center of the sphere.

Table 3 shows the calculation of charging free energies for single sodium and chloride ions in spherical clusters of water molecules. As can be seen in this table, the charging free energies are approximately converged at around 10 Å.[41] Furthermore, except for one outlier, the multigrid algorithm provides results close to the no-cutoff limit. In comparison, Lee and Warshel[8] have shown that simple radial truncation of electrostatics without an approximate long-range treatment can lead to charging free energy errors greater than 10 kcal/mol.

Next, we performed 2-ns molecular dynamics simulations of protein L in four different environments: a layer without the GB reaction field, a hybrid layer, a hybrid sphere, and a PME cube. Figure 7 shows that the root-mean-square deviation (RMSD) trajectories from the original X-ray structure are similar between the different types of simulations. Both of the layer calculations seem to wander less from the X-ray structure than the sphere and PME simulations. The layer w/o GB simulation is especially restricted throughout most of the 2 ns. The RMSD trajectory results are consistent with the *B*-factor data presented in Figure 8 and Table 4. As can be seen in Table 4, the average *B*-factor calculated from the layer w/o GB simulation is noticeably smaller than the PME calculation, while the sphere simulation is very close. The reduced fluctuations of the layer simulations may be due in part to the angular restraint imposed on the moment of inertia of the solute [eq. (14)]. Another possibility is that water molecules in the layer may have restricted flow leading to slight constraints on the flexibility of the solute. Perhaps, the layer w/o GB simulation is restricted more than the hybrid layer calculation because charge–charge interactions within the solute are not sufficiently shielded. Although there is good agreement among the four protocols for calculating *B*-factors, none of the simulation methods produces *B*-values in quantitative agreement with experimental data. This discrepancy can be attributed to a variety of issues, such as a lack of anisotropic components in the fluctuation calculations, crystal packing effects, and forcefield errors.

In Figure 9, the distances between the α-carbons of the final simulation structures and the X-ray structure are presented. One can see that the hybrid layer simulation has larger deviations than the other simulation methods in the main helical region (residues 23–34) of the polypeptide. Upon inspection of the structure, the helix in the layer simulation is still intact but is shifted slightly relative to the rest of the protein. One possible reason for this result is that the helix dipole is interacting with the artificial dipole that exists at the boundary of the simulation volume.

Finally, the wall clock times for the four different protein simulations are presented in Table 5. As one can see, the layer protocol is almost three times faster than the PME method for this protein. However, this is due in large part to the fact that the number of water molecules has been reduced by a factor of 4.5. The computational times of the layer and sphere methods suggest worse than linear scaling versus the number of water molecules. This result is likely due to the fact that the shape of the simulation volume affects the prefactor of scaling. One reason why the hybrid algorithm is slower than PME on a per atom basis is because the hybrid method must compute two pairwise energy terms: electrostatic *and* GB.

Our single snapshot error analysis results are in good agreement with the work of Skeel et al.[29] after one accounts for the fact that, in contrast to their protocol, we have used standard intramolecular nonbonded exclusion lists. These exclusions remove, for example, the large intramolecular O–H electrostatic interactions of the TIP3 water molecule, which can make percentage errors appear smaller by roughly an order of magnitude. The results of Skeel et al.[29] suggest that the multigrid method does not achieve as high a degree of accuracy as fast multipole methods. The accuracy in the multigrid approach is mainly improved through extension of the short-range cutoff,[29] which can quickly become computationally inefficient because the short-range atom pair list increases as the cube of the cutoff radius. Fast multipole methods have been shown

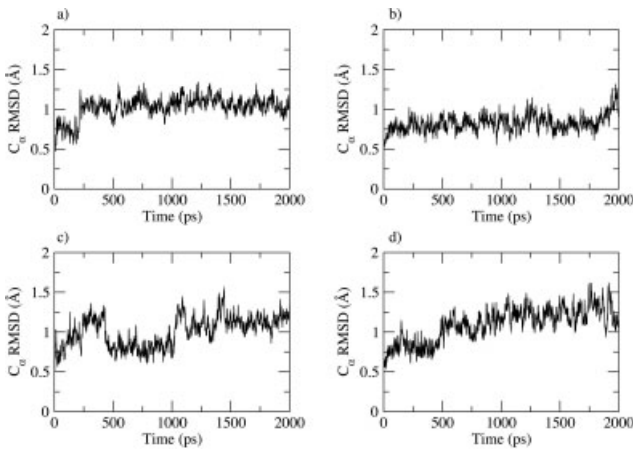

Figure 7. Trajectories of the C_α RMSD from the X-ray structure for the three types of protein L simulations. (a) Layer, (b) layer (w/o GB), (c) sphere, (d) PME.

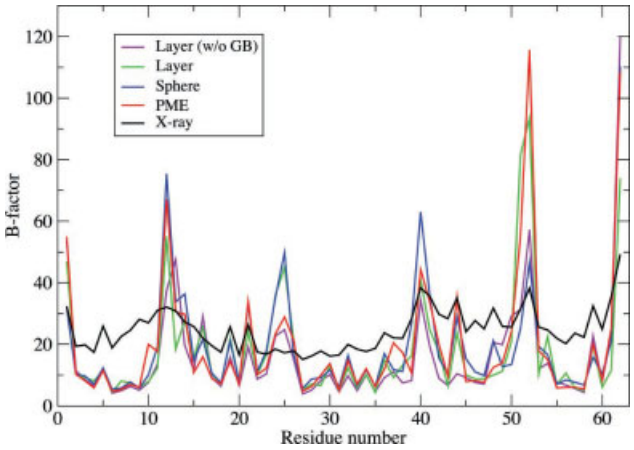

Figure 8. Calculated *B*-factors (Å^2) of each residue for the three types of protein L simulations compared to the experimental values. Calculated values were obtained from the second half of the 2-ns simulations.

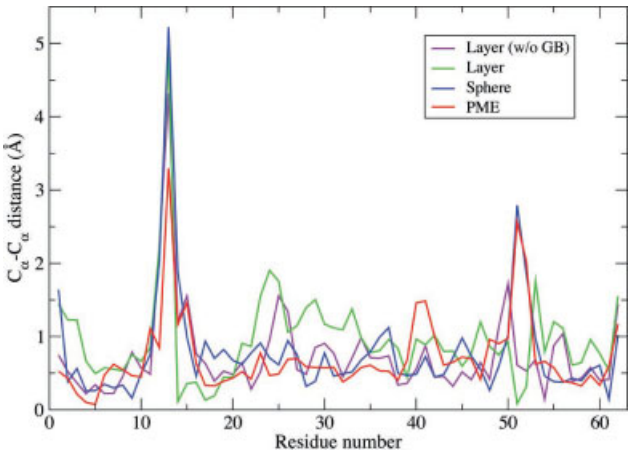

Figure 9. C_α distances to the X-ray structure for the final structures generated with the four types of protein L simulations.

to provide highly accurate long-range treatments, although some difficulties exist in the evaluation of forces unless high precision tolerances are employed. The multigrid method has an advantage over these multipole methods because it utilizes a smoothly varying potential that leads to relatively stable molecular dynamics trajectories. This smoothness feature is also useful if multiple time stepping is desired.[29]

An alternative approach to the treatment of the long-range electrostatic interactions is the local reaction field (LRF) method of Warshel and coworkers.[8] The LRF method builds a very accurate long-range treatment at specified time intervals in a molecular dynamics calculation. It then uses the same long-range potential to rapidly update the long-range energy and forces within each interval. It is not obvious how the energies and forces deviate compared to the exact limit as the system evolves from the initial long-range potential update step. However, free energy studies using the LRF technique have provided accurate results vs. exact treatments.[7,8] Unlike the LRF method, the full multigrid quantity in our procedure is updated at every time step, such that there are no assumptions regarding the smoothness of the long-range potential over time.[8] Instead, the multigrid approach induces a smoothness assumption in the spatial dimensions that likely impairs its ability to achieve high accuracy. In any case, the most crucial aspect of the multigrid approach for the purpose of this work is that it is the only current method that can easily be adapted to handle the complex GB potential term.

Although the initial results presented do not comprise a comprehensive analysis of the hybrid method compared to established methods, such as PME, we would like to distinguish the types of applications where the method may be appropriate and computationally efficient. First, the hybrid approach would be useful for simulating very large solutes where the box size required to encapsulate the solute in water molecules is prohibitive.[11] An important caveat is that the layer method may not be reliable for protein simulations where major conformation rearrangements are expected to occur, as the simulation volume is held fixed. For simulations of large flexible movements of molecules to be feasible, one would need to make intelligent additions of volume at the beginning of the simulation or multiple restarts of the simulation as the solute conformation changes. Furthermore, the fixed layer approach would be counter-productive for folding/unfolding studies.[45] The layer method may also exaggerate protein stability somewhat as was seen in the results presented in this work. In contrast, the spherical cluster approach should work fine for cases where large rearrangements are expected to occur. The sphere method is about twice as slow as the layer approach for the protein studied in this work. This may be an acceptable trade-off given that

Table 4. Average Per-Residue *B*-Factors for the Last 1 ns of Each of the Four Different Types of MD Simulations of Protein L.

Method	Average *B*-factor (Å^2)
Layer (w/o GB)	16.1
Layer	18.2
Sphere	19.3
PME	19.7
X-ray	24.7

For simulation details, see Methods section.

Table 5. Wall Clock Times for the Four Different Types of MD Simulations of Protein L.

Method	Time (days)	Number of water molecules
Layer (w/o GB)	3.4	1378
Layer	7.2	1378
Sphere	13.2	3091
PME	19.6	6231

For simulation details, see the Methods section.

the spherical approach has somewhat better agreement with PME. However, for certain biological systems that are highly nonspherical (e.g., a DNA strand), the layer approach would likely be more appropriate since the computational savings would be more substantial.[10]

Finite cluster and non-Ewald PBC approaches are often preferred for certain applications where complications regarding real–virtual interactions of the solute may be an issue.[46,7] For example, in the calculation of charging free energies of polyatomic solutes in solution, certain corrections to the Ewald formulation must be employed to remove real–virtual interactions.[47] Other applications where complications might arise include the calculation of mutational free energies where net charge changes occur and the computation of pK_as of ionizable residues. Many large-scale charging calculations to date have been performed with non-Ewald PBC.[48,49] Unfortunately, non-Ewald PBC methods require relatively large cutoffs, and hence, increased computational effort for the nonbonded electrostatic terms vs. PME. The hybrid method, on the other hand, thanks to the multigrid enhancement and reduction of explicit water molecules, achieves similar computational efficiencies as PME without incurring the artificial periodicity of the solute. In fact, in a separate article,[41] we applied the hybrid approach towards calculating solvation free energies of fixed protein conformations. The resultant data are then used to benchmark the accuracy of various Poisson implicit solvent models.

Similar to other finite-cluster methods, such as GSBP and SCAAS, the hybrid method is not limited to explicit solvation of the entire solute.[20,50] This is a major computational advantage vs. PBC approaches. For example, in certain applications, such as protein–ligand binding, protein–protein binding, and mutational studies, the simulation volume can be defined in terms of a subset of "active" atoms. The nonactive atoms can then be treated with a regular GB model. Because the simulation volume is fixed in the current procedure, the nonactive atoms should be fixed also.[20]

In principle, the hybrid method can be extended to allow for a dynamically adjusting layer that can achieve constant pressure.[10,11] There are, however, several difficulties with a constant pressure approach that may need to be addressed. First, large conformational changes of the solute may lead to significant changes in the geometric volume of the simulation volume. Thus, it may be necessary to add or delete water molecules using a constant chemical potential approach such as a grand canonical Monte Carlo procedure.[51] Moreover, a dynamic layer adjusting to local pressures may squeeze out waters from unfavorable locations near a hydrophobic region of the solute.[11] This problem suggests that the layer widths for each heavy atom may have to be restrained from becoming too small. In addition, coupling of the boundary to the solute positions may lead to decreased fluctuations of the solute atoms due to the fact that waters near the boundary will be dragged around as the solute moves. Finally, the simplicity and computational efficiency of using a Born radii grid to interpolate atomic Born radii will have to be sacrificed if the boundary needs to be updated often.

Conclusion

We have introduced a new algorithm for performing biomolecular simulations of a solute surrounded by an irregularly shaped layer of water molecules. Unlike previous layer-based approaches, proper reaction field treatment is achieved by encapsulating the explicit simulation volume in an implicit solvent described by the GB theory. Significant computational enhancement of this approach is achieved through the use of a pairwise multigrid technique that has been extended to incorporate the GB model. This method overall can be faster than the conventional PME technique in cases where the number of water molecules can be significantly reduced. We foresee the application of the hybrid technique in simulations where a finite cluster of water molecules is preferred, for example, the calculation of binding affinities of protein–ligand and protein–protein complexes,[52] and the computation of mutational free energies.[53]

Acknowledgments

Computational time for this work was provided in part by the U.S. Army Research Laboratory Major Shared Resource Center and the Wake Forest University Distributed Environment for Academic Computing cluster. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Army.

References

1. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Oxford University Press: New York, 1987.
2. King, G.; Warshel, A. J Chem Phys 1989, 91, 3647.
3. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. J Am Chem Soc 1990, 112, 6127.
4. Bursulaya, B. D.; Brooks, C. L., III. J Phys Chem B 2000, 104, 12378.
5. Calimet, N.; Schaefer, M.; Simonson, T. Proteins Struct Funct Genet 2001, 45, 144.
6. Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. Proteins Struct Funct Biol 2004, 55, 502.
7. Sham, Y. Y.; Warshel, A. J Chem Phys 1998, 109, 7940.
8. Lee, F. S.; Warshel, A. J Chem Phys 1992, 97, 3100.
9. Masunov, A.; Lazaridis, T. J Am Chem Soc 2003, 125, 1722.
10. Beglov, D.; Roux, B. J Chem Phys 1994, 100, 9050.
11. Lounnas, V.; Ludemann, S. K.; Wade, R. C. Biophys Chem 1999, 78, 157.
12. Alper, H.; Levy, R. M. J Chem Phys 1993, 99, 9847.
13. Neumann, M.; Steinhauser, O. Mol Phys 1980, 39, 437.
14. van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. J Chem Phys 1998, 108, 10220.
15. Brooks, C. L., III; Karplus, M. J Chem Phys 1983, 79, 6312.
16. Brooks, C. L., III; Brünger, A.; Karplus, M. Biopolymers 1985, 24, 843.
17. Darden, T.; Pearlman, D.; Pedersen, L. G. J Chem Phys 1998, 109, 10921.
18. Essex, J. W.; Jorgensen, W. L. J Comp Chem 1995, 16, 951.
19. Kirkwood, J. G. Chem Rev 1936, 19, 275.
20. Im, W.; Berneche, S.; Roux, B. J Chem Phys 2001, 114, 2924
21. Beglov, D.; Roux, B. Biopolymers 1995, 35, 171.

22. Rosenhouse–Dantsker, A.; Osman, R. Biophys J 2000, 79, 66.
23. Kentsis, A.; Mezei, M.; Osman, R. Biophys J 2003, 84, 805.
24. Lee, M. S.; Salsbury, F. R., Jr.; Brooks, C. L., III. J Chem Phys 2002, 116, 10606
25. Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L., III. J Comp Chem 2003, 24, 1348.
26. Stote, R. H.; States, D. J.; Karplus, M. J Chim Phys 1991, 88, 2419.
27. Greengard, L.; Rokhlin, V. J Comput Phys 1987, 73, 325.
28. Sandak, B. J Comp Chem 2001, 22, 717.
29. Skeel, R. D.; Tezcan, I.; Hardy, D. J. J Comp Chem 2002, 23, 673.
30. Brooks, C. L., III; Karplus, M. J Mol Biol 1989, 208, 159.
31. Warshel, A. Chem Phys Lett 1978, 55, 454.
32. Warshel, A. J Phys Chem 1979, 83, 1640.
33. Im, W.; Beglov, D.; Roux, B. Comp Phys Commun 1998, 111, 59.
34. Im, W.; Lee, M. S.; Brooks, C. L., III. J Comp Chem 2003, 24, 1691.
35. Onufriev, A.; Case, D. A.; Bashford, D. J Comp Chem 2002, 23, 1297.
36. Salsbury, F. R., Jr.; Lee, M. S.; Feig, M.; Brooks, C. L., III, to appear.
37. Lawson, C. L. SIAM Rev 1965, 7, 415.
38. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminatham, S.; Karplus, M. J Comp Chem 1983, 4, 187.
39. Mackerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph–McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz–Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
40. Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327.
41. Lee, M. S.; Olson, M. A. J Phys Chem B, submitted.
42. O'Neill, J. W.; Kim, D. E.; Baker, D.; Zhang, K. Y. J. Acta Crystallogr D 2001, 57, 480.
43. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. J Chem Phys 1995, 103, 8577.
44. Herce, D. H.; Darden, T.; Sagui, C. J Chem Phys 2003, 119, 7621.
45. Sheinerman, F. B.; Brooks, C. L. I. J Mol Biol 1998, 278, 439.
46. Sham, Y. Y.; Chu, Z. T.; Warshel, A. J Phys Chem B 1997, 101, 4458-4472.
47. Bogusz, S.; Cheatham, T. E. I.; Brooks, B. R. J Chem Phys 1998, 108, 7070.
48. Zhang, L. Y.; Gallichio, E.; Friesner, R. A.; Levy, R. M. J Comp Chem 2001, 22, 591.
49. Bürgi, R.; Kollman, P. A.; van Gunsteren, W. F. Proteins 2002, 47, 469.
50. Lee, F. S.; Chu, Z. T.; Warshel, A. J Comp Chem 1993, 14, 161.
51. Adams, D. J. Mol Phys 1975, 29, 307.
52. Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. Proteins Struct Funct Gene 2000, 39, 393.
53. Huo, S.; Massova, I.; Kollman, P. A. J Comp Chem 2002, 23, 15.
54. Hoogstraten, C. G.; Choe, S.; Westler, W. M.; Markley, J. L. Protein Sci 1995, 4, 2289.
55. Kurinov, I. V.; Harrison, R. W. Nat Struct Biol 1994, 1, 735.
56. Kirkwood, J. G. In Theory of Liquids; Alder, B. J., Ed.; Gordon and Breach: New York, 1968.